

FIG. 2

DOCUMENT ID	URL	TITLE	REFERENCED DOCUMENT	LINK IMPORTANCE
00001	http://www.fujitsu.co.jp/	FUJITSU HOME		1023
00002	http://www.kantei.go.jp/	OFFICIAL RESIDENCE OF PRIME MINISTER		2055
⋮				

DOCUMENT INFORMATION TABLE 41

DOCUMENT ID	URL SIMILARITY
00006	3
00138	2
⋮	

REFERENCED DOCUMENT TABLE 42

FIG. 3

KWD ID	REPRESENTATIVE WORD	OCCURRENCE DOCUMENT
001	コンピュータ (konpyu-ta: COMPUTER)	
....		

KEYWORD TABLE 51

KEYWORD	PRONUNCIATION CHARACTERS	KWD ID
computer	COMPUTER	001
コンピュータ	こんびゆた	001
プログラム	ぶろぐらむ	002
⋮		

KEYWORD RELATION TABLE 52

DOCUMENT ID	OCCURRENCE
00001	8
00003	1
⋮	

OCCURRENCE DOCUMENT TABLE 53

FIG. 4

CHARACTER STRING	FOLLOWING CHARACTER STRINGS	KEYWORD STRINGS
TOP	あ (a), い (i), ..	
あ (a)	あいぼ (aibo), あお (ao), ...	
あいぼ (aibo)		相棒 (aibou:MATE), アイボリー (aiborī : IVORY).
あお (ao)	あおぞ (aozora)	青 (ao: BLUE), 蒼 (ao: DARK BLUE), ...

INDEX INFORMATION TABLE 61

KEYWORD ID	CORRELATED DOCUMENT ID STRINGS
093 321	0005, 0008, .. 0004, 0008, ...

CORRELATED DOCUMENT TABLE 62

DOCUMENT ID	CORRELATED KEYWORD ID STRINGS
0005 0008 . .	093, 099, 122, ... 093, 156, 321, ...

CORRELATED KEYWORD TABLE 63

FIG. 5

FORMAT IN yyyyymmddH-MM

DATE AND TIME	KWD ID	DOCUMENT ID
200001121436	003	00123
200001121437	005	00054
.		
.		
.		

ACCESS LOG 71

FIG. 6

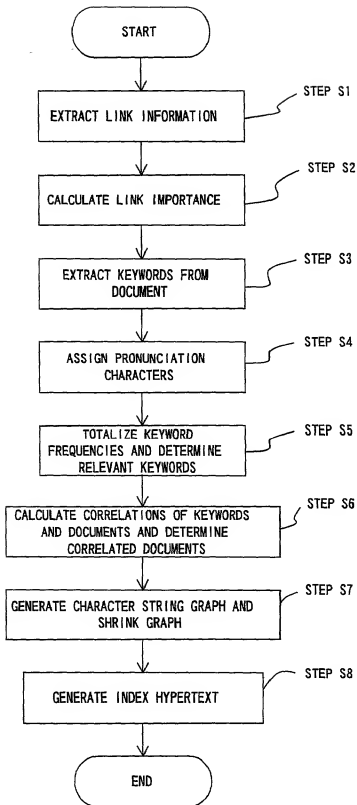
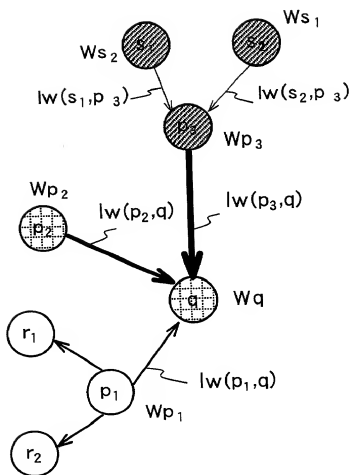


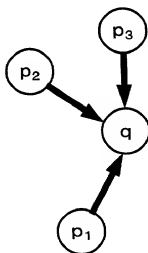
FIG. 7



CIRCLE(O) : WEB PAGE
 THICKNESS OF ARROW(\rightarrow) : LINK WEIGHT
 PATTERN OF CIRCLE(O) : URL SIMILARITY

FIG. 8

$$\text{sim}(p_i, q) = 1$$

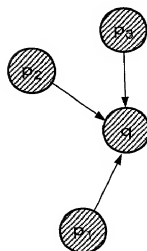


$$lw(p_i, q) = \frac{1}{\text{sim}(p_i, q)} = 1$$

$$w_q = C_q + w_{p1} + w_{p2} + w_{p3}$$

FIG. 9A

$$\text{sim}(p_i, q) = n + 1$$



$$lw(p_i, q) = \frac{1}{\text{sim}(p_i, q)} = \frac{1}{n + 1}$$

$$w_q = C_q + \frac{w_{p1} + w_{p2} + w_{p3}}{n + 1}$$

FIG. 9B

NUMBER OF DOMAINS CONTAINED IN HIGHER RANKED URLS

NUMBER OF DOMAINS

NUMBER OF HIGHER RANKED URLS (%100)

SEQUENCE 1
SEQUENCE 2
SEQUENCE 3
SEQUENCE 4
SEQUENCE 5

FIG. 10

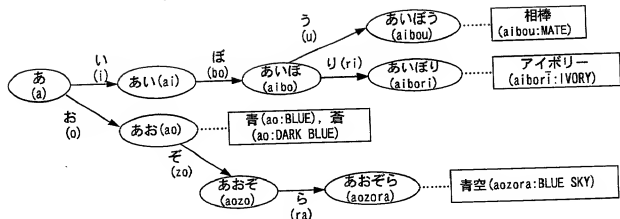


FIG. 11A

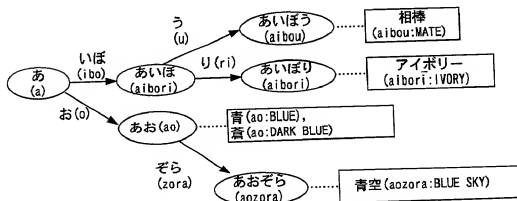


FIG. 11B

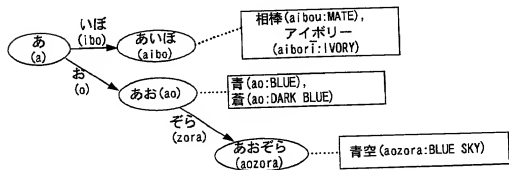


FIG. 11C

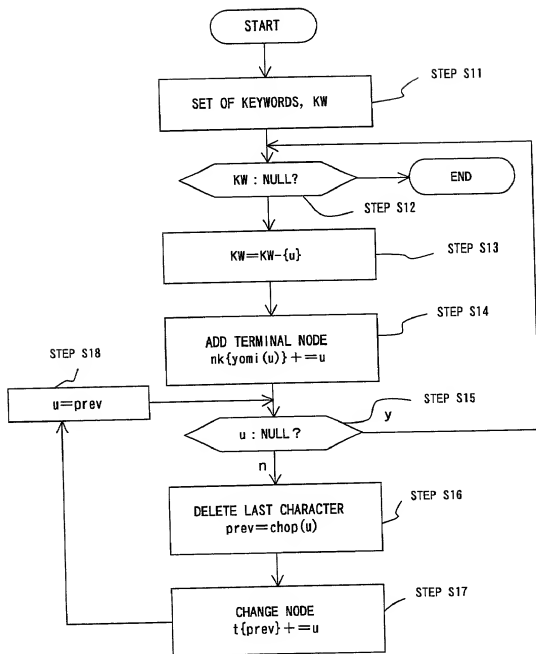


FIG. 12

```

proc init_kw_graph (
{
  @KW:set of keywords;      # SET OF KEYWORDS
  yomi : YOMI/Spell of keywords; # FUNCTION OR ARRAY THAT RETURNS PRONUNCIATION CHARACTERS OF KEYWORD
  foreach u in KW {         # FOR EACH KEYWORD u
    nk{yomi{u}} .= u.+";    # DESIGNATE nk() OF NODE OF PRONUNCIATION CHARACTERS OF KEYWORD u
    for ( i=0; i<length(u); i++) { # REPEAT FOR LENGTH OF CHARACTER STRING OF KEYWORD u
      local prev = chop(u);      # DELETE LAST CHARACTER OF KEYWORD u AND ADD TO PARENT NODE
      t{prev} .= u.+";
      u = prev;
    }
  }
}
)

```

FIG. 13

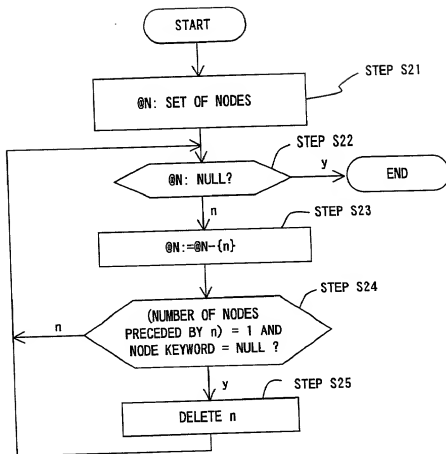


FIG. 14

```

proc shrink_middle ()
{
  @N : set of nodes
  foreach n (@N) {
    next = t{n};    # NEXT NODE LIST
    kw = nk{n};     # KEYWORD LIST
    if (length(next) == 1 && kw == "") {
      delete(n)    # DELETE NODE n
    }
  }
}

```

FIG. 15

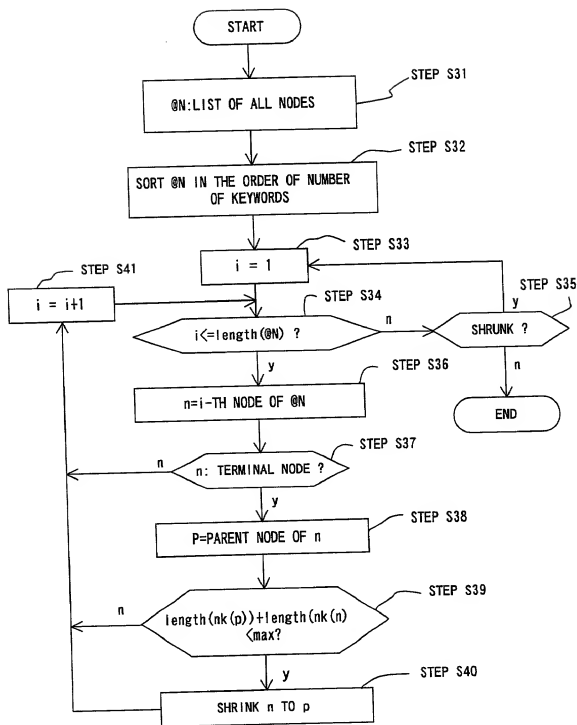


FIG. 16


```

proc shrink_leaf ()
{
    @N: set of nodes;      # NODE LIST
    word_max = 2;          # word_max : IN THIS EXAMPLE, 2
    changed = true;        # # WHEN KEYWORD IS TRANSFERRED, true
    @N = sort by _nk_length @N; # SORTING IN ASCENDING ORDER OF NUMBER OF KEYWORDS
    while (changed) {      # CONTINUING WHILE TRANSFER IS PERFORMED
        changed = false;
        foreach n in @N {
            if (is_leaf(n)) { # IN THE CASE OF TERMINAL NODE
                p = parent_node(n); # PARENT NODE
                if (length(nk{p}) + length(nk{n}) < word_max) {
                    nk{p} = nk{n}."+"; # TRANSFERRING KEYWORD
                    delete (n); # DELETE TERMINAL NODE
                    changed = true; # PROOF OF TRANSFER
                }
            }
        }
    }
}

```

FIG. 17

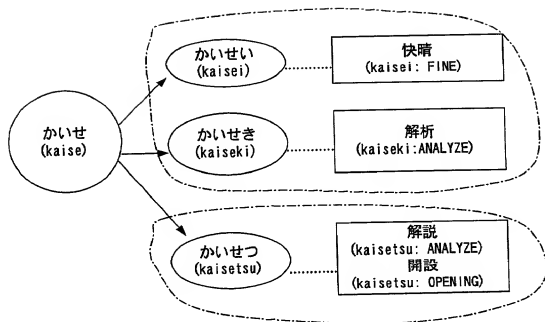


FIG. 18

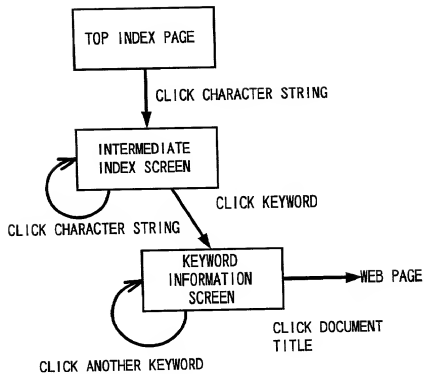


FIG. 19

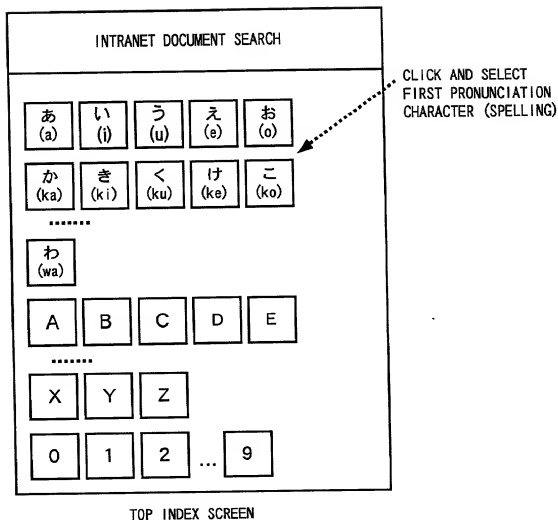


FIG. 20

INTELLECTUAL 50-KANA CHARACTER INDEX OF INTRA-COMPANY PAGES

あ(a)	い(i)	う(u)	え(e)	お(o)	が(ga)	ぎ(gi)	ぐ(gu)	げ(ge)	ご(go)
か(ka)	き(ki)	く(ku)	け(ke)	こ(ko)	ざ(za)	じ(ji)	ず(zu)	ぜ(ze)	ぞ(zo)
さ(sa)	し(shi)	す(su)	せ(se)	そ(so)	だ(da)	ぢ(de)	ど(do)		
た(ta)	ち(chi)	つ(tsu)	て(te)	と(to)	は(ha)	ひ(bi)	ぶ(bu)	べ(be)	ぼ(bo)
な(na)	に(ni)	ぬ(nu)	ね(ne)	の(no)					
は(ha)	ひ(bi)	ふ(fu)	へ(he)	ほ(ho)					
	び(bi)	ぶ(bu)	べ(he)	ほ(ho)					
ま(ma)	み(mi)	む(mu)	め(me)	も(mo)					
や(ya)	ゆ(yu)	よ(yo)							
ら(ra)	り(ri)	る(ru)	れ(re)	ろ(ro)					
わ(wa)									
A B C D E F G H I J K L M N O P Q R S T U V W X 1 2 3 4 5									

(NOTE) " " LONG SOUND SHOULD BE REMOVED. SELECT " " (tu) AND " " (ya) FOR " " (tu) AND " " (ya).

CLEAR

SEARCH FOR A KEYWORD INCLUDING

FIG. 21

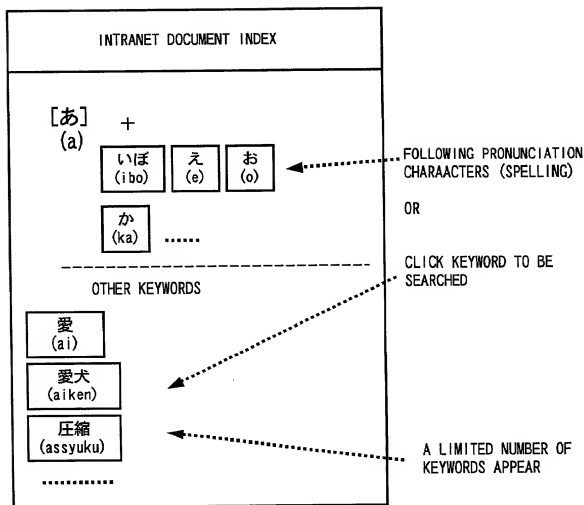


FIG. 22

"い" (i)		
<p>い (i) え (e) ろ (ro)</p> <p>か (ka) き (ki) く (ku) け (ke) こ (ko) う (u) が (ga) ぎ (gi) ご (go)</p> <p>さ (sa) し (shi) す (su) せ (se) し (ji)</p> <p>た (ta) ち (chi) づ (tsu) で (de) ん (un) ぶ (bu)</p> <p>に (ni) の (no) べ (be) ん (un) と (to)</p> <p>み (mi) め (me) じ (ji)</p> <p>や (ya) ゆ (yu) よ (yo)</p> <p>ら (ra) り (ri) り (ru) れ (re)</p> <p>わ (wa) ん (un)</p>	<p>(NOTE) "ー" LONG SOUND SHOULD BE REMOVED. SELECT "っ (tu)" AND "ゃ (ya)" FOR "っ (tu)" AND "ゃ (ya)".</p>	<p>OTHER KEYWORDS</p> <ul style="list-style-type: none"> ・イオン (ion) ・イネーブル (inēburu) ・伊豆 (izu) ・位相 (isou) ・依存 (izon) ・依存性 (izonnsei) ・意図 (ito) ・意欲 (iyoku) ・移転 (itenn) ・違反 (ihann) ・違反行為 (ihannkoui) ・遺族 (izoku) ・医薬 (iyaku) ・医用 (iyou) ・稲城長沼 (inaginaganuma) ・急ぎ (isogi) ・居酒屋 (izakaya) ・指宿 (ibusuki) ・色刺激 (iroshigeki) ・田舎 (inaka)

FIG. 23

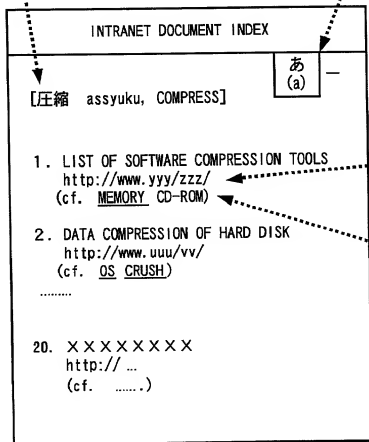
"いべんと" (ibento)

KEYWORDS

- ibento
- ibento ID
- ibento karendā
- ibento sabisu
- ibento sukeijyūru
- ibento maikurosofuto
- ibento rogu
- ibento annai
- ibento kudou
- ibento tetsuzuki
- ibento tsuchi
- ibento naiyou
- ibento hassei
- ibento horyu
- kaisaiyotei ibento
- kanushi ibento
- yotei ibento

REPRESENTATIVE WORD AND SYNONYM

ROUTE PATH



WHEN DOCUMENT
INFORMATION IS
CLICKED, JUMPED TO
RELEVANT PAGE

WHEN CORRELATED
KEYWORD IS CLICKED,
JUMPED TO KEYWORD
INFORMATION SCREEN
OF THE KEYWORD

F I G. 25

トップ (toppu) - イベント (ibento)

「イベントカレンダー」

(IBENTO KARENDĀ : EVENT CALENDAR)

MAJOR PAGES ABOUT "イベントカレンダー"

- <http://www.paso.co.jp/event/2000.html> (03/17/1999)
2000 NEN KARENDĀ : CALENDAR OF YEAR 2000
(KEYWORDS : ソフトウェア (software), 展覧会 (exhibition))
- <http://www.cal.co.jp/event9907.html> (06/23/1999)
7 GATSU NO MOYOUSHI : EVENT ON JULY
(KEYWORDS : 音楽会 (ongakukai), コンサート (konsāto : concert))
- <http://www.yahoo.co.jp/event/> (06/23/1999)
イベントリスト (ibento risto : EVENT LIST)

FIG. 26

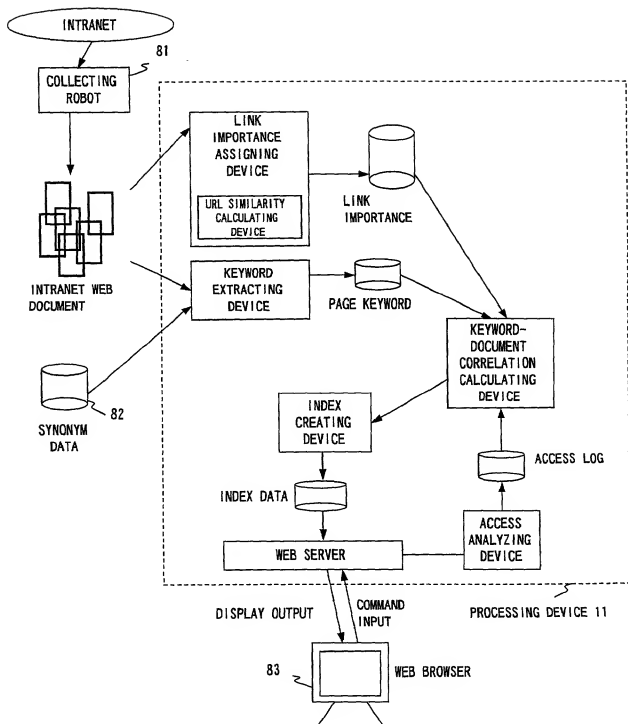


FIG. 27

The diagram illustrates a system for processing in-Internet web documents. The system is divided into a **PROCESSING DEVICE 11** (indicated by a dashed line) and a **WEB BROWSER**.

Components and Data Flow:

- INTRANET** (oval) connects to a **COLLECTING ROBOT** (rectangle).
- The **COLLECTING ROBOT** collects **INTRANET WEB DOCUMENT** (represented by overlapping rectangles).
- SYNONYM DATA** (cylinder) is input to the **KEYWORD EXTRACTING DEVICE** (rectangle).
- The **INTRANET WEB DOCUMENT** is processed by the **LINK IMPORTANCE ASSIGNING DEVICE** (rectangle) and the **URL SIMILARITY CALCULATING DEVICE** (rectangle).
- The **LINK IMPORTANCE ASSIGNING DEVICE** outputs **LINK IMPORTANCE** (cylinder).
- The **URL SIMILARITY CALCULATING DEVICE** outputs **PAGE KEYWORD** (cylinder).
- The **KEYWORD EXTRACTING DEVICE** outputs **PAGE KEYWORD** (cylinder).
- The **LINK IMPORTANCE**, **PAGE KEYWORD**, and **DOCUMENT TYPE** (cylinder, output from **DOCUMENT TYPE DETERMINING DEVICE 91**) are inputs to the **KEYWORD-DOCUMENT CORRELATION CALCULATING DEVICE** (rectangle).
- The **KEYWORD-DOCUMENT CORRELATION CALCULATING DEVICE** outputs to the **INDEX CREATING DEVICE** (rectangle).
- The **INDEX CREATING DEVICE** outputs **INDEX DATA** (cylinder).
- The **INDEX DATA** is stored in the **WEB SERVER** (rectangle).
- The **WEB SERVER** provides **DISPLAY OUTPUT** to the **WEB BROWSER** (monitor icon).
- The **WEB BROWSER** sends **COMMAND INPUT** to the **WEB SERVER**.
- The **WEB SERVER** also sends data to the **DOCUMENT TYPE DETERMINING DEVICE 91** (rectangle).
- The **DOCUMENT TYPE DETERMINING DEVICE 91** outputs **DOCUMENT TYPE** (cylinder).
- The **ACCESS LOG** (cylinder) is input to the **ACCESS ANALYZING DEVICE** (rectangle).
- The **ACCESS ANALYZING DEVICE** outputs to the **KEYWORD-DOCUMENT CORRELATION CALCULATING DEVICE**.

F I G. 28

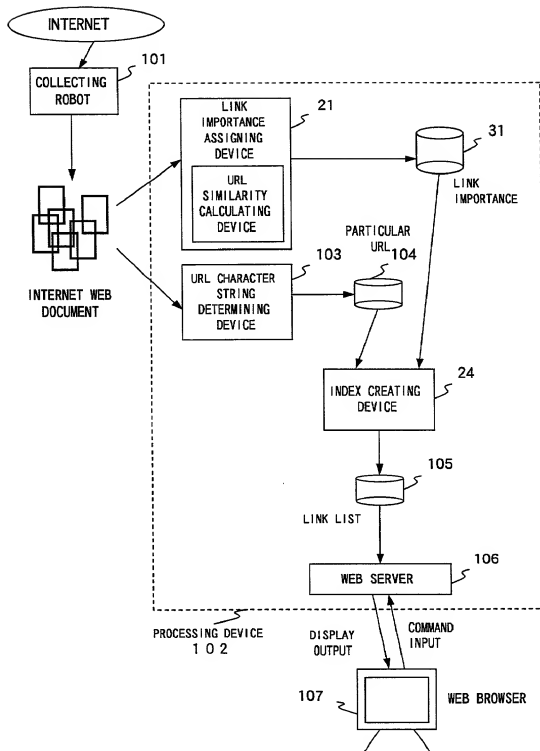


FIG. 29

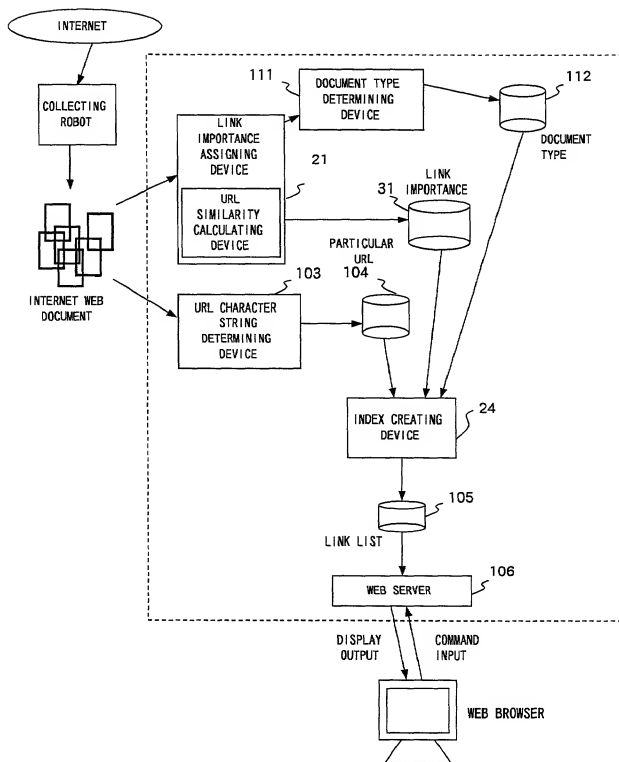


FIG. 30

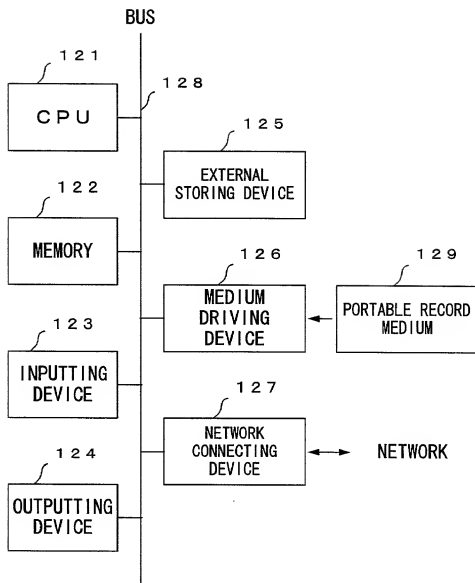


FIG. 31

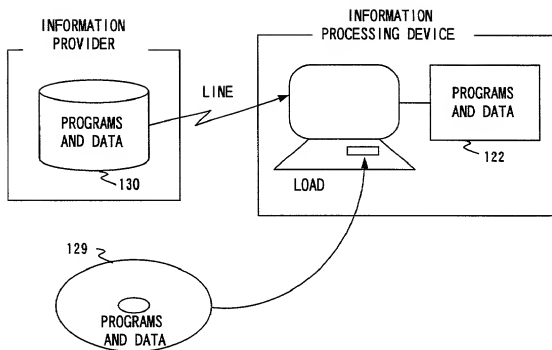


FIG. 32